Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 7 (2010) 130–134

Procedia Engineering

www.elsevier.com/locate/procedia

2010 Symposium on Security Detection and Information Processing

Partial Least-squares Regression for Identification of Liquid Materials

Wei Li [a,b], Yu Zhong [a], Yu Zhang [a,b], Daoyang Yu [a], Bai Sun [a], Minqiang Li [a,*], Jinhuai Liu [a,*]

[a] *Key Laboratory of Biomimetric Sensing and Advanced Robot Technology, Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, 230031, PR China*

[b] *Department of Automation, University of Science and Technology of China, Hefei, 230031, PR China*

Abstract

Partial least-squares (PLS) regression was demonstrated as a promising suitable method to identify quantitatively liquid materials based on the profile of energy dispersive X-ray scattering spectrum in this paper. Methanol, ethanol, sulfuric acid and acetic acid of different concentrations were measured under conditions of the scatter angle of 7° and the count time of 5 min. The results showed that the regression of concentration was effective. This suggests that PLS can be employed in detecting liquid materials.

Keywords: partial least-squares; energy dispersive X-ray scattering; liquid materials; liquid explosives; precursor chemicals.

1. Introduction

Liquid materials such as hazardous liquid fuel, liquid explosives and precursor chemicals are paid increasingly attention. Many different approaches have been proposed to identify liquid materials, including Raman scattering [1], gas chromatography/mass spectrometry [2], magnetic resonance [3-4], terahertz spectroscopy [5], Hilbert spectroscopy [6] and ion mobility spectrometry [7]. In contrast, the energy dispersive X-ray scattering (EDXRS) has been proven to be a suitable one with its specific advantages such as non-destructive, non-invasive, low-cost and high-resolution [8-12].

It is well-known that liquids do not possess sharp Bragg peaks which are generally presented by crystalline specimen. Under the same experiment conditions, each liquid exhibits a unique scattering profile, including the position, shape, width and intensity of a series of peaks which are governed by the liquid properties, and those can be regarded as the 'fingerprint' of the detected liquid. Because the diffraction profile，which is composed by multichannel signal, can be seen as multivariable data, PLS, which is good at processing high-dimensional and inter-correlated data is usually used [8,9,12].

In this paper, we introduced PLS regression to identify quantitatively liquid materials using EDXRS. Methanol, ethanol, sulfuric acid and acetic acid were chosen as the representation of liquid materials.

* Corresponding author. Tel.: +86 0551 5591167; fax: +86 0551 5592420.
E-mail address: jhliu@iim.ac.cn, mqli@iim.ac.cn.

1877-7058
doi:10.1016/j.proeng.2010.11.019

2. The experimental system and methods

2.1. The experimental system

The atoms in crystalline materials are arranged in such a way that regularly spaced atomic planes can be identified. The relationship between the X-ray energy E, the scatter angle θ and an atomic planar spacing (lattice spacing) d of the investigated material is given by the Bragg's law [13]:

$$2d\sin\frac{\theta}{2}=\frac{nhc}{E}$$

where h is Planck's constant, c is velocity of light. A primary X-ray photon with high energy is scattered under a small scatter angle θ for a given lattice spacing d. Under fixed scatter angle, different lattice spacings characterizing the crystalline structure of the material result in different diffraction peaks within the energy spectrum of the scattered X-ray photons.

The EDXRS spectrometer of our experiment is shown in figure 1, which consists of a polychromatic source of X-ray and an energy-resolving detector [14]. The polychromatic source of X-ray comprises an X-ray tube and the related collimators to define the incident and scattered beams. The W target is adopted in the range 0-40 kV and 0-30 mA. A liquid-nitrogen-cooled high purity germanium (HP Ge) detector with sufficient energy resolution is used. The Canberra planar HP Ge detector with 50 mm^2 detector area and 5 mm thick crystal has a resolution of 140 eV at 5.9 keV combined with a Canberra spectrum master InSpector 2000.

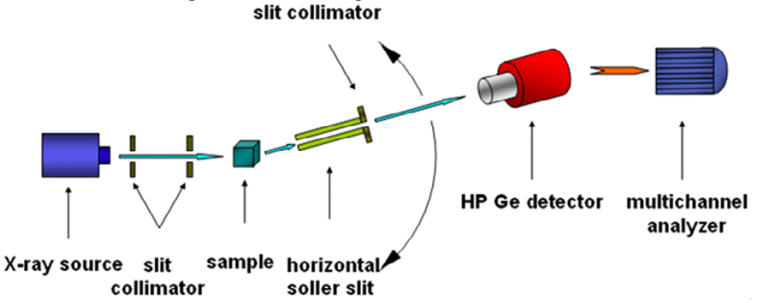

Fig. 1. The geometry of the scattering system.

2.2. The principle of partial least-squares

PLS is a method for constructing predictive models based on principal component analysis (PCA) when the factors are many and highly collinear [15]. The linear regression model is:

$$y_i=\beta_0+B'_{q,p}\,x_i+e_i \tag{1}$$

PLS assumes that the x- and y-variables are related through a bilinear model:

$$x_i=\bar{x}+P_{p,k}\,\tilde{t}_i+g_i$$

$$y_i=\bar{y}+A'_{q,k}\,\tilde{t}_i+f_i$$

In this model, $\bar{x}$ and $\bar{y}$ represent the mean of the x- and the y-variables. The $\tilde{t}_i$ is called the scores which are k-dimensional, with k << p, whereas $P_{p,k}$ is the matrix of x-loadings. The residuals of each equation are represented by the g_i and f_i respectively. The matrix $A_{q,k}$ denotes the slope matrix in the regression of y_i on $\tilde{t}_i$.

The estimates for the parameters in the original model (1) are:

$$\hat{B}_{p,q}=R_{p,k}\,\hat{A}_{k,q}$$

$$\hat{\beta}_0 = \bar{y} - \hat{B}'_{q,p}\,\bar{x}\,.$$

Due to the combination of PCA and multiple linear regressions, PLS is capable for highly collinear data. The tool used to perform PLS was the MATLAB® library for Robust Analysis (LIBRA) [16].

3. Results and discussions

Methanol, ethanol, sulfuric acid and acetic acid of different concentrations (10% to 90%) were measured four times under conditions of the scatter angle of 7° and the count time of 5 min. Figure 2 showed the result of methanol in the first measurement. The spectrum consisted of 1024 points in *x*-coordinate. Peaks near 150 are useless due to the target peak, while ones after 200 are useful. In order to focus the key feature of tested materials and reduce the computational complexity, we selected data between 201 and 712 in *x*-coordinate.

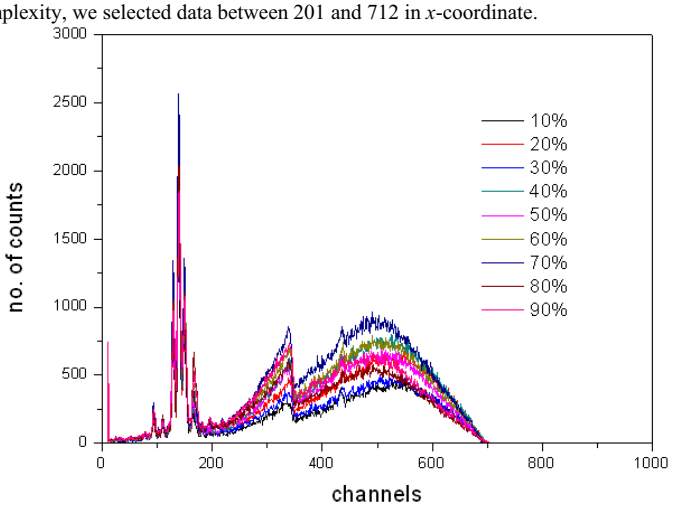

Fig. 2. EDXRS spectrum of methanol of different concentrations.

The data of each liquid formed a matrix with 36 rows and 512 columns, then they were split into two sets: one was training set, used to build and calibrate a multivariate model, and the other was prediction set, used to predict the concentration. According to PLS, the scattering profile is X and the concentration is Y. All data were analyzed by LIBRA.

The prediction results of four kinds of liquid were shown in Figure 3-6. The trend of predicted concentration with actual concentration was obvious. However, there were fluctuations in some concentrations. The reason may be the container of liquid. In our experiment, the samples were concealed in plastic bags, whose shapes were changed easily in each experiment. When it occurred, the optical path was changed, which led to the error of prediction. In the further, standard containers will be adopted to optimize the experiment.

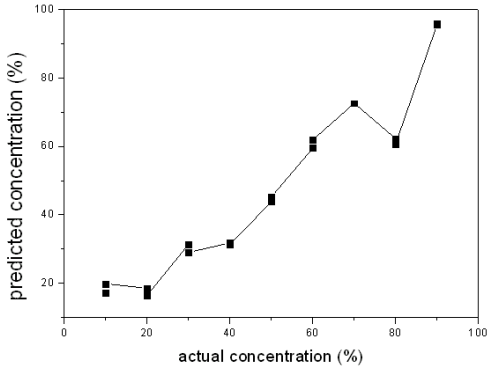

Fig. 3. The prediction of methanol.

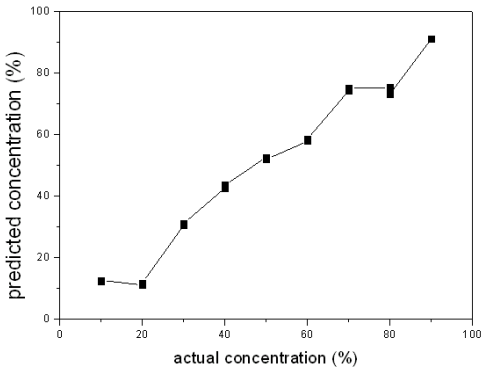

Fig. 4. The prediction of ethanol.

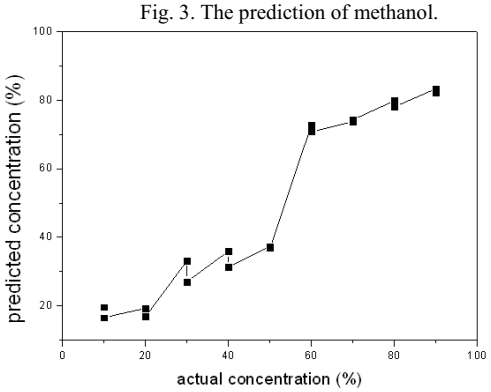

Fig. 5. The prediction of sulfuric acid.

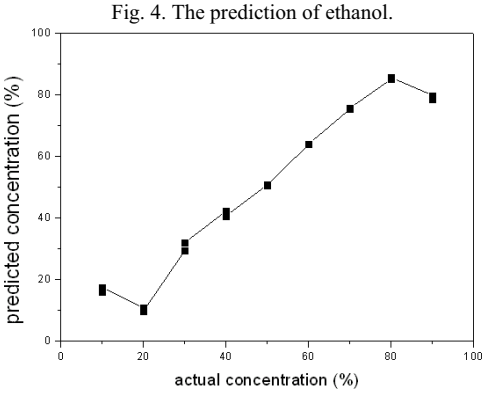

Fig. 6. The prediction of acetic acid.

4. Conclusion

In this paper, the feasibility of applying PLS to identify quantitatively liquid materials has been demonstrated. Methanol, ethanol, sulfuric acid and acetic acid of different concentrations were measured under conditions of the scatter angle of 7° and the count time of 5 min by using EDXRS. The results of PLS regression indicated that the trend of predicted concentration with actual concentration was obvious. Our findings have shown a new opportunity for developing effective detecting approach towards liquid materials.

Acknowledgement

This work was financially supported by the Key Program of National Natural Science Foundation of China (Grant 10635070).

References

[1] Eliasson C, Macleod NA, Matousek P. *Anal. Chem*. 2007; **79**: 8185.
[2] Lu Y, Harrington PB. *Anal. Chem.* 2007; **79**: 6752.
[3] Mauler J, Danieli E, Casanova F, Blumich B. *NATO Sci. Peace Secur. Ser. B Phys. Biophys.* 2009; 193.
[4] Kumar S, Prado PJ. *NATO Sci. Peace Secur. Ser. B Phys. Biophys.* 2009; 73.
[5] Ikeda T, Matsushita A, Tatsuno M et al.. *Appl. Phys. Lett.* 2005; **87**: 034105.
[6] Divin Y, Poppe U, Urban K. *NATO Sci. Peace Secur. Ser. B Phys. Biophys.* 2008; 189.

[7] Burks RM, Hage DS. *Anal. Bioanal. Chem.* 2009; **395**: 301.
[8] Luggar RD, Farquharson MJ, Horrocks JA et al.. *X-Ray Spectrom.* 1998; **27**: 87.
[9] Cook EJ, Fong R, Horrocks JA et al.. *Appl. Radiat. Isot*. 2007; **65**: 959.
[10] Kämpfe B, Luczak F, Michel B. *Part. Part. Syst. Charact.* 2005; **22**: 391.
[11] Cook EJ, Griffiths JA, Koutalonis M et al.. *Proceeding of Spie*, 2009; 7310, 73100I.
[12] Cook EJ, Pani S, George L et al.. *IEEE Trans. Nucl. Sci.* 2009; **56**: 1459.
[13] Johns PC, Yaffe MJ. Med. *Phys*. 1983; **10**: 40.
[14] Sun B, Li M, Zhang F et al.. *Microchem J*. 2010; **95**: 293.
[15] De Jong S. Chemometrics Intell. *Lab. Syst.* 1993; **18**: 251.
[16] Verboven S and Hubert M. Chemometrics Intell. *Lab. Syst.* 2005; **75**: 127.